

0420  
0590  
1009

OIPE

A 2

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/829,382

DATE: 04/26/2001

TIME: 17:00:05

Input Set : N:\Crf3\RULE60\09829382.txt

Output Set: N:\CRF3\04262001\I829382.raw

## SEQUENCE LISTING

ENTERED

## 4 (1) GENERAL INFORMATION:

6 (i) APPLICANT: Masure, H. Robert  
7 Rosenow, Carsten I.  
8 Tuomanen, Elaine  
9 Wizemann, Theresa M.

11 (ii) TITLE OF INVENTION: CHOLINE BINDING PROTEINS FOR  
12 ANTI-PNEUMOCOCCAL VACCINES

14 (iii) NUMBER OF SEQUENCES: 25

16 (iv) CORRESPONDENCE ADDRESS:

17 (A) ADDRESSEE: David A. Jackson, Esq.  
18 (B) STREET: 411 Hackensack Ave, Continental Plaza, 4th  
19 Floor  
20 (C) CITY: Hackensack  
21 (D) STATE: New Jersey  
22 (E) COUNTRY: USA  
23 (F) ZIP: 07601

25 (v) COMPUTER READABLE FORM:

26 (A) MEDIUM TYPE: Floppy disk  
27 (B) COMPUTER: IBM PC compatible  
28 (C) OPERATING SYSTEM: PC-DOS/MS-DOS  
29 (D) SOFTWARE: PatentIn Release #1.0, Version #1.30

31 (vi) CURRENT APPLICATION DATA:

C--> 32 (A) APPLICATION NUMBER: US/09/829,382  
C--> 33 (B) FILING DATE: 09-Apr-2001  
34 (C) CLASSIFICATION:

36 (vii) PRIOR APPLICATION DATA:

37 (A) APPLICATION NUMBER: 08/847,065  
38 (B) FILING DATE:

41 (viii) ATTORNEY/AGENT INFORMATION:

42 (A) NAME: Jackson Esq., David A.  
43 (B) REGISTRATION NUMBER: 26,742  
44 (C) REFERENCE/DOCKET NUMBER: 600-1-158 ..

46 (ix) TELECOMMUNICATION INFORMATION:

47 (A) TELEPHONE: 201-487-5800  
48 (B) TELEFAX: 201-343-1684

51 (2) INFORMATION FOR SEQ ID NO: 1:

53 (i) SEQUENCE CHARACTERISTICS:

54 (A) LENGTH: 30 amino acids  
55 (B) TYPE: amino acid  
56 (C) STRANDEDNESS:  
57 (D) TOPOLOGY: linear

59 (ii) MOLECULE TYPE: peptide

60 (A) DESCRIPTION: cBP112

62 (iii) HYPOTHETICAL: NO

64 (iv) ANTI-SENSE: NO

66 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

RAW SEQUENCE LISTING                      DATE: 04/26/2001  
 PATENT APPLICATION: US/09/829,382        TIME: 17:00:05

Input Set : N:\Crf3\RULE60\09829382.txt  
 Output Set: N:\CRF3\04262001\I829382.raw

```

W--> 68      Xaa Glu Asn Glu Gly Ser Thr Gln Ala Ala Thr Ser Ser Asn Met Ala
69          1              5              10              15
71      Lys Thr Glu His Arg Lys Ala Ala Lys Gln Val Val Asp Glu
72          20              25              30
75 (2) INFORMATION FOR SEQ ID NO: 2:
77      (i) SEQUENCE CHARACTERISTICS:
78          (A) LENGTH: 10 amino acids
79          (B) TYPE: amino acid
80          (C) STRANDEDNESS:
81          (D) TOPOLOGY: linear
83      (ii) MOLECULE TYPE: peptide
84          (A) DESCRIPTION: cBP90
86      (iii) HYPOTHETICAL: NO
88      (iv) ANTI-SENSE: NO
91      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
93      Ala Arg Glu Phe Ser Leu Glu Lys Thr Arg
94          1              5              10
97 (2) INFORMATION FOR SEQ ID NO: 3:
99      (i) SEQUENCE CHARACTERISTICS:
100          (A) LENGTH: 23 amino acids
101          (B) TYPE: amino acid
102          (C) STRANDEDNESS:
103          (D) TOPOLOGY: linear
105      (ii) MOLECULE TYPE: peptide
106          (A) DESCRIPTION: cBP84
108      (iii) HYPOTHETICAL: NO
110      (iv) ANTI-SENSE: NO
113      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
W--> 115     Xaa Arg Glu Phe Ser Leu Glu Lys Thr Arg Asn Ile Gly Ile Met Ala
116         1              5              10              15
118         His Val Asp Ala Gly Lys Thr
119             20
122 (2) INFORMATION FOR SEQ ID NO: 4:
124      (i) SEQUENCE CHARACTERISTICS:
125          (A) LENGTH: 27 amino acids
126          (B) TYPE: amino acid
127          (C) STRANDEDNESS:
128          (D) TOPOLOGY: linear
130      (ii) MOLECULE TYPE: peptide
131          (A) DESCRIPTION: cBP80
133      (iii) HYPOTHETICAL: NO
135      (iv) ANTI-SENSE: NO
138      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
W--> 140     Xaa Lys Xaa Xaa Trp Gln Xaa Lys Gln Tyr Leu Lys Glu Asp Gly Ser
141         1              5              10              15
W--> 143     Gln Ala Ala Asn Glu Xaa Val Phe Asp Thr Ala
144             20              25
147 (2) INFORMATION FOR SEQ ID NO: 5:
149      (i) SEQUENCE CHARACTERISTICS:

```

## RAW SEQUENCE LISTING

DATE: 04/26/2001

PATENT APPLICATION: US/09/829,382

TIME: 17:00:06

Input Set : N:\Crf3\RULE60\09829382.txt

Output Set: N:\CRF3\04262001\I829382.raw

```

150      (A) LENGTH: 30 amino acids
151      (B) TYPE: amino acid
152      (C) STRANDEDNESS:
153      (D) TOPOLOGY: linear
155      (ii) MOLECULE TYPE: peptide
156      (A) DESCRIPTION: cBP78
158      (iii) HYPOTHETICAL: NO
160      (iv) ANTI-SENSE: NO
163      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:
165      Gln Lys Ile Ile Gly Ile Asp Leu Gly Thr Thr Asn Ser Ala Val Ala
166      1          5          10          15
168      Val Leu Glu Gly Thr Glu Ser Lys Ile Ile Ala Asn Pro Glu
169      20          25          30
172 (2) INFORMATION FOR SEQ ID NO: 6:
174      (i) SEQUENCE CHARACTERISTICS:
175      (A) LENGTH: 16 amino acids
176      (B) TYPE: amino acid
177      (C) STRANDEDNESS:
178      (D) TOPOLOGY: linear
180      (ii) MOLECULE TYPE: peptide
181      (A) DESCRIPTION: cBP70
183      (iii) HYPOTHETICAL: NO
185      (iv) ANTI-SENSE: NO
188      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:
W--> 190      Xaa Xaa Xaa Glu Val Ala Lys Xaa Ser Gln Asp Thr Thr Thr Ala Ser
191      1          5          10          15
194 (2) INFORMATION FOR SEQ ID NO: 7:
196      (i) SEQUENCE CHARACTERISTICS:
197      (A) LENGTH: 19 amino acids
198      (B) TYPE: amino acid
199      (C) STRANDEDNESS:
200      (D) TOPOLOGY: linear
202      (ii) MOLECULE TYPE: peptide
203      (A) DESCRIPTION: cBP60
205      (iii) HYPOTHETICAL: NO
207      (iv) ANTI-SENSE: NO
209      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:
W--> 211      Xaa Asn Glu Arg Val Lys Ile Val Ala Thr Leu Gly Pro Ala Val Glu
212      1          5          10          15
214      Gly Arg Gly
217 (2) INFORMATION FOR SEQ ID NO: 8:
219      (i) SEQUENCE CHARACTERISTICS:
220      (A) LENGTH: 18 amino acids
221      (B) TYPE: amino acid
222      (C) STRANDEDNESS:
223      (D) TOPOLOGY: linear
225      (ii) MOLECULE TYPE: peptide
226      (A) DESCRIPTION: cBP50 pep
228      (iii) HYPOTHETICAL: NO

```

## RAW SEQUENCE LISTING

DATE: 04/26/2001

PATENT APPLICATION: US/09/829,382

TIME: 17:00:06

Input Set: N:\Crif3\RULE60\09829382.txt

Output Set: N:\CRF3\04262001\I829382.raw

```

230      (iv) ANTI-SENSE: NO
233      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:
W--> 235      Xaa Ile Ile Xaa Xaa Val Tyr Ala Arg Glu Val Leu Asp Ser Arg Gly
236          1             5             10             15
238      Asn Pro
241 (2) INFORMATION FOR SEQ ID NO: 9:
243      (i) SEQUENCE CHARACTERISTICS:
244          (A) LENGTH: 12 amino acids
245          (B) TYPE: amino acid
246          (C) STRANDEDNESS:
247          (D) TOPOLOGY: linear
249      (ii) MOLECULE TYPE: peptide
250          (A) DESCRIPTION: cBP112-Int1
252      (iii) HYPOTHETICAL: NO
254      (iv) ANTI-SENSE: NO
257      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:
259      Glu Asp Arg Arg Asn Tyr Pro Thr Asn Thr Tyr Lys
260          1             5             10
263 (2) INFORMATION FOR SEQ ID NO: 10:
265      (i) SEQUENCE CHARACTERISTICS:
266          (A) LENGTH: 11 amino acids
267          (B) TYPE: amino acid
268          (C) STRANDEDNESS:
269          (D) TOPOLOGY: linear
271      (ii) MOLECULE TYPE: peptide
272          (A) DESCRIPTION: cBP112-Int2
274      (iii) HYPOTHETICAL: NO
276      (iv) ANTI-SENSE: NO
279      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:
W--> 281      Xaa Asp Asp Gln Gln Ala Glu Glu Asp Tyr Ala
282          1             5             10
285 (2) INFORMATION FOR SEQ ID NO: 11:
287      (i) SEQUENCE CHARACTERISTICS:
288          (A) LENGTH: 8 amino acids
289          (B) TYPE: amino acid
290          (C) STRANDEDNESS:
291          (D) TOPOLOGY: linear
293      (ii) MOLECULE TYPE: peptide
295      (iii) HYPOTHETICAL: NO
297      (iv) ANTI-SENSE: NO
300      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:
302      Trp Gln Pro Pro Arg Ala Arg Ile
303          1             5
305 (2) INFORMATION FOR SEQ ID NO: 12:
307      (i) SEQUENCE CHARACTERISTICS:
308          (A) LENGTH: 16 amino acids
309          (B) TYPE: amino acid
310          (C) STRANDEDNESS:
311          (D) TOPOLOGY: linear

```

RAW SEQUENCE LISTING                      DATE: 04/26/2001  
 PATENT APPLICATION: US/09/829,382        TIME: 17:00:06

Input Set : N:\Crif3\RULE60\09829382.txt  
 Output Set: N:\CRF3\04262001\I829382.raw

```

313      (ii) MOLECULE TYPE: peptide
314      (A) DESCRIPTION: 50 KDa protein
321      (iii) HYPOTHETICAL: NO
322      (iv) ANTI-SENSE: NO
316      (v) FRAGMENT TYPE: N-terminal
318      (vi) ORIGINAL SOURCE:
319          (A) ORGANISM: S. pneumoniae
325      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:
327      Ile Thr Asp Val Tyr Ala Arg Glu Val Leu Asp Ser Arg Gly Asn Pro
328      1          5          10          15
331 (2) INFORMATION FOR SEQ ID NO: 13:
333      (i) SEQUENCE CHARACTERISTICS:
334          (A) LENGTH: 16 amino acids
335          (B) TYPE: amino acid
336          (C) STRANDEDNESS:
337          (D) TOPOLOGY: linear
339      (ii) MOLECULE TYPE: peptide
340      (A) DESCRIPTION: 50 KDa protein
342      (iii) HYPOTHETICAL: NO
344      (iv) ANTI-SENSE: NO
346      (v) FRAGMENT TYPE: N-terminal
348      (vi) ORIGINAL SOURCE:
349          (A) ORGANISM: B. subtilis
352      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:
354      Ile Val Asp Val Arg Ala Arg Glu Val Leu Asp Ser Arg Gly Asn Pro
355      1          5          10          15
358 (2) INFORMATION FOR SEQ ID NO: 14:
360      (i) SEQUENCE CHARACTERISTICS:
361          (A) LENGTH: 210 base pairs
362          (B) TYPE: nucleic acid
363          (C) STRANDEDNESS: double
364          (D) TOPOLOGY: linear
366      (ii) MOLECULE TYPE: cDNA
368      (iii) HYPOTHETICAL: NO
370      (iv) ANTI-SENSE: NO
372      (vi) ORIGINAL SOURCE:
373          (A) ORGANISM: S. pneumoniae
376      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:
378 TCCCGCGGCA ACCCAACT TGAAGTAGAA GTTTACTACTG AATCAGGTGC TTTCGGACGT      60
380 GGTATGGTTC CATCAGGAGC TTCTACTGGT GAACACGAAG CAGTTGAACT TCGCGACGGT      120
382 GACAAATCTC GTTACGGTGG TCTTGGTACA CAAAAAGCTG TTGACAACGT AAACAACATC      180
384 ATTGCTGAGG CCATCATGG CTACGATGTA      210
387 (2) INFORMATION FOR SEQ ID NO: 15:
389      (i) SEQUENCE CHARACTERISTICS:
390          (A) LENGTH: 210 base pairs
391          (B) TYPE: nucleic acid
392          (C) STRANDEDNESS: double
393          (D) TOPOLOGY: linear
395      (ii) MOLECULE TYPE: cDNA

```

## VERIFICATION SUMMARY

PATENT APPLICATION: US/09/829,382

DATE: 04/26/2001

TIME: 17:00:07

Input Set : N:\Crf3\RULE60\09829382.txt

Output Set: N:\CRF3\04262001\I829382.raw

L:32 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]  
L:33 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]  
L:68 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1  
L:115 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3  
L:140 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4  
L:143 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4  
L:190 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:6  
L:211 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7  
L:235 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8  
L:281 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:10